Research Note 2013-01

Formulating the Brogden Classification Framework as a Discrete Choice Model

Tirso E. Diaz
Human Resources Research Organization

November 2012

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved for distribution:

MICHELLE SAMS, Ph.D.
Director

Research accomplished under contract
for the Department of the Army by

Human Resources Research Organization

Technical review by

Peter Greenston, U.S. Army Research Institute

NOTICES

DISTRIBUTION: This Research Note has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: Destroy this Research Note when it is no longer needed. Do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Research Note are not to be construed as an official Department of the Army position, unless so designated by other authorized document.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy)	2. REPORT TYPE	3. DATES COVERED (from. . . to)
November 2012	Final	September 2011 – February 2012

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Formulating the Brogden Classification Framework as a Discrete Choice Model	W5J9CQ-11-C-0045
	5b. PROGRAM ELEMENT NUMBER 622785
6. AUTHOR(S) Tirso E. Diaz	5c. PROJECT NUMBER A790
	5d. TASK NUMBER 329
	5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Human Resources Research Organization 66 Canal Center Plaza, Suite 700 Alexandria, Virginia 22314	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. MONITOR ACRONYM
U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg 1464 / Mail Stop: 5610) Fort Belvoir, VA 22060-5610	ARI
	11. MONITOR REPORT NUMBER Research Note 2013-01

12. DISTRIBUTION/AVAILABILITY STATEMENT Distribution Statement A: Approved for public release, distribution is unlimited.

13. SUPPLEMENTARY NOTES

Contracting Officer's Representative and Subject Matter Expert POC: Dr. Peter Greenston

14. ABSTRACT *(Maximum 200 words)*:

The Brogden optimal classification framework measures potential classification benefits of predictors by assigning applicants to the jobs that will maximize predicted performance subject to job quota constraints. Current implementations of Brogden's framework do not include classification policy constraints (e.g., cut scores and gender restriction), applicant preferences, or the impact of other classification tools available to the Army (e.g., monetary incentives to channel applicants to particular job training). To accommodate elements of real world classification systems and thereby better inform operational problems, this research reformulated Brogden's classification framework using discrete choice modeling. We specified a mixed multinomial logit model for classification that is mathematically equivalent to a multivariate normal based implementation of Brogden's framework. We also proposed an empirical or sample based method for classification analysis based on the multinomial logit (MNL) model that can accommodate personnel classification policy constraints, such as cut scores and gender restriction, and is robust to the functional form and distribution of the criterion estimates. Illustrative example applications of the MNL classification model showed the expected effects of policy constraints, with practical implications for the analysis results.

15. SUBJECT TERMS

Selection and classification; Brogden optimal classification; Discrete choice models; Job choice

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	47	Dorothy E. Young 703-545-2316

Research Note 2013-01

Formulating the Brogden Classification Framework as a Discrete Choice Model

Tirso E. Diaz
Human Resources Research Organization

Personnel Assessment Research Unit
Tonia S. Heffner, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
6000 6th Street, Bldg 1464
Fort Belvoir, VA 22060

November 2012

Army Project Number
622785A790

Performance and
Training Technology

FORMULATING THE BROGDEN CLASSIFICATION FRAMEWORK AS A DISCRETE CHOICE MODEL

EXECUTIVE SUMMARY

Research Requirement:

In the 1950's Brogden proposed a number of related approaches for measuring the classification efficiency of a predictor battery for assigning applicants to different jobs (Brogden, 1954; Brogden, 1955; Brogden, 1959). These approaches are all based on optimal classification in which applicants are assigned to the jobs for which they have the highest predicted performance. More recent implementations of Brogden's framework have been developed for computing the mean predicted performance (MPP) under different statistical distributional assumptions (De Corte 2000; Chen and Darby 1997). These implementations, however, do not address classification policy constraints (e.g., cut scores and gender restriction), applicant preferences, or the impact of other recruiting/classification tools available to the Army (e.g., monetary incentives that channel applicants to particular job training). To obtain policy guidance it is important to reformulate Brogden's optimal classification framework as a statistical model that can inform operational classification problems.

A discrete choice model (DCM) is a rich modeling framework that is concerned with agents making choices among alternatives. In the Army classification situation, the discrete choice process can describe applicants choosing among job training opportunities. The elements in Brogden's work are comparable to those in a DCM; for example, assigning an applicant to the job with the highest predicted performance corresponds to an individual choosing an alternative with maximum utility. The goals of this effort are: (a) to demonstrate the feasibility of formulating Brogden's optimal classification problem as a DCM; (b) to show how a DCM can accommodate classification policy constraints, such as cut scores and gender restrictions; and (c) to explore how a DCM could be applied in an expanded classification framework that takes into account applicant preferences.

Procedure:

To show that the Brogden classification framework can be formulated using a DCM framework, we rescaled the augmented criterion estimates proposed by Brogden using an arbitrarily large constant and then added independent standard Gumbel errors to obtain the mathematical structure of utility equations in a DCM. Using this transformation, we derived a mixed multinomial logit (MMNL) classification model that is analytically comparable to de Corte's (2000) implementation of Brogden's model based on an assumed multivariate normal (MVN) distribution for criterion estimates. We also derived an empirical or sample based multinomial logit (MNL) classification model that can accommodate personnel classification policy constraints and is robust with regard to the distributional assumption for criterion estimates. We illustrated the MNL optimal classification model by evaluating the classification

efficiency potential of the Armed Services Vocational Aptitude Battery (ASVAB) and the Tailored Adaptive Personality Assessment Screen (TAPAS) for minimizing 6-month attrition in four military occupational specialties (MOS) under pure classification and selection-classification models and varying classification policy requirements. Lastly, we described a hybrid approach for evaluating classification potential of predictors that account for applicant preferences.

Findings:

Brogden's optimal classification framework for evaluating potential classification benefits of predictors can be solved using DCM methods. The population based MMNL classification model was shown to be equivalent to the MVN-based optimal classification model of de Corte. The empirical or sample based MNL classification model provides a rich framework that can accommodate personnel classification policy constraints, such as cut scores and gender restriction, and is robust to the functional form and distribution of the criterion estimates. In illustrative examples, while changes in average estimated attrition due to inclusion of policy constraints were negligible, they were consistent with anticipated effects and are expected to become stronger with more valid criterion models and greater differentiation in eligibility constraints among jobs. The possibility of using criterion estimates that are not normally distributed was shown to have practical implications for the analysis results.

Utilization and Dissemination of Findings:

The DCM methods developed in this report provide alternative approaches for evaluating potential classification benefits of predictors that are numerically convenient and can accommodate classification policy constraints. For classification problems where sufficient applicant data are available, the empirical MNL classification model can more accurately evaluate potential classification efficiency compared to population-based classification models (i.e., de Corte's approach and MMNL classification models).

FORMULATING THE BROGDEN CLASSIFICATION FRAMEWORK AS A DISCRETE CHOICE MODEL

CONTENTS

Page

Page

LIST OF TABLES

FORMULATING THE BROGDEN CLASSIFICATION FRAMEWORK AS A DISCRETE CHOICE MODEL

Research Requirement

In the 1950's Brogden proposed a number of related approaches for measuring the classification efficiency of a predictor battery for assigning applicants to different jobs (Brogden, 1954; Brogden, 1955; Brogden, 1959). These approaches are all based on optimal classification in which applicants are assigned to the jobs for which they have the highest predicted performance. Earlier efforts by Brogden and other researchers (e.g., Alley & Darby 1994; Chen and Darby 1997) for measuring the classification efficiency assumed equal predictor validities, equal predictor inter-correlations, and equal job quotas. Advances in mathematical computing have made it possible to apply Brogden's general approach for measuring classification efficiency with unequal predictor validities and inter-correlation among the predictors (e.g., De Corte 2000) – for example, to evaluate the incremental classification efficiency of non-cognitive experimental predictor batteries beyond the ASVAB.

While current implementations of Brogden's framework may be sufficient for evaluating the potential incremental classification efficiency of competing experimental predictor batteries, they do not offer guidance for formulating operational classification policies. These implementations do not address classification policy constraints (e.g., cut scores and gender restriction), applicant preferences, or the impact of other recruiting/classification tools available to the Army (e.g., monetary incentives that channel applicants to particular job training). It is not surprising, therefore, that realized classification efficiency is often much lower than levels suggested by optimal MPP.

To obtain policy guidance it is necessary to reformulate Brogden's optimal classification framework as a statistical model that can inform operational classification problems. In this regard, a promising direction is to formulate Brogden's framework as a discrete choice model (DCM). Choice analysis is concerned with how agents make choices among alternatives, where some of the factors are observable to the researcher and others are not (and are treated stochastically). In the Army classification situation, the discrete choice process can describe applicants choosing among job training opportunities. The elements in Brogden's work are comparable to those in a DCM; for example, assigning an applicant to the job with the highest predicted performance corresponds to an individual choosing an alternative with maximum utility. Beyond algorithmic similarity, a DCM would provide researchers and policy makers (when evaluating or formulating classification policy tools) with a statistical framework that can accommodate important factors in real world classification.

The goals of this effort are: (a) to demonstrate the feasibility of formulating the Brogden optimal classification framework as a DCM; (b) to show how a DCM can accommodate classification policy constraints, such as cut scores and gender restrictions; and (c) to explore how a DCM could be applied in an expanded classification framework that takes into account applicant preferences as a function of monetary incentives and cognitive / non-cognitive predictors.

Background

Brogden Classification Framework

Brogden (1959) proposed to measure classification efficiency of a predictor battery using the maximum obtainable average predicted performance across all possible ways of assigning *n* applicants to *m* jobs. This maximum allocation average is also known as the MPP (mean predicted performance) and the allocation that produced it is said to be optimal.

In applications, the predicted performance scores of a given applicant for the *m* jobs are often assumed to be jointly distributed as multivariate normal (MVN). Under the special case with equal predictor validities, equal predictor inter-correlations, and equal quotas across jobs, optimal classification is equivalent to assigning applicants to the jobs corresponding to their highest predicted performance scores. Using this heuristic, the MPP can be computed by averaging the highest among the *m* predicted performance scores of applicants. In the general case, with unequal predictor validities, or unequal predictor inter-correlations, or unequal target quotas, this simple strategy does not work.

Brogden (1955) proposed an assignment strategy based on the augmented predicted criterion (or augmented criterion estimate) of applicants, in which job specific constants are added to the *m* criterion estimates of applicants. Each applicant is then assigned to the job corresponding to his highest augmented criterion estimate. Brogden showed that by appropriately choosing the job specific constants, assigning applicants to jobs in this manner satisfies the target job quotas and produces the maximum average predicted performance scores (i.e., the MPP).

Applying Brogden's augmented criterion estimate assignment strategy involves non-trivial analytic and numerical computing challenges. Earlier efforts by Brogden and other researchers (e.g., Alley & Darby 1994; Chen and Darby 1997) for measuring the classification efficiency (i.e. the MPP) of a predictor battery focused on the special case with equal predictor validities, equal predictor inter-correlations, and equal quotas. An exception is De Corte (2000) who developed an approach for the general optimal classification problem by analytically working directly with the multivariate normal joint distribution of the criterion scores. De Corte's approach will be used in this research to represent the general Brogden framework and to examine its relationship to DCM methodology. Earlier studies used de Corte's approach to evaluate the optimal classification of experimental predictors in the Army (Ingerick, Diaz, & Putka, 2009).

DCM Job Choice Model Applications

In contrast to classification analysis, in which individual applicants are classified to jobs to provide maximum benefit for the organization (i.e., maximize MPP), choice modeling focuses on identifying jobs that are expected to provide maximum benefit for individual applicants (i.e., maximize utility). A Job Choice Model (JCM) is concerned with describing actual applicant job choices in relation to their characteristics and the attributes of job alternatives under a random utility maximization (RUM) assumption. RUM posits that among all available alternatives, an

individual will choose the alternative that has maximum value to him/her. In other words, a JCM describes optimal jobs for an applicant while a classification model describes optimal jobs from the viewpoint of personnel managers.

Earlier Army personnel classification studies used various forms of the JCM to model applicant MOS choices based on data about the specific military occupational specialty (MOS) and incentives that were presented to applicants. Diaz, Ingerick, and Sticha (2007a) used a JCM based on a nested logit model to simulate applicant's choice of MOS to support the implementation of an unobtrusive, simulation-based evaluation of the Enlisted Personnel Allocation System (EPAS; Sticha, Diaz, Greenston, & McWhite, 2007).

In a later project, Diaz, Ingerick, and Sticha (2007b) extended the model to consider prediction of MOS-term of service (TOS) combinations, and applied to analysis of the impact of increasing the individual cap on recruiting bonuses. In response to a difficult recruiting environment, the Army obtained legislative authority to increase the EB program from $20K to $40K. The increased incentives could expand the recruiting market and channel applicants from other MOS into ones with higher incentives. The main focus of the study was to estimate the channeling effects of expanded alternative bonus programs.

More recently, Diaz, Sticha, Hogan, Mackin, and Greenston (2012) estimated a JCM that models Army applicants' MOS and TOS enlistment preferences as a function of enlistment incentives to support the development of a tool that can assist the Enlistment Incentive Review Board (EIRB) allocate incentives to the MOS and TOS options that provide the greatest incremental benefit to the Army. They implemented the analysis capability of the JCM as a proof-of-concept Decision Support Tool (DST) that allows users to specify incentive policy scenarios, predict applicant enlistments by MOS and TOS and associated costs for each policy scenario, and compare the results across different policy scenarios.

While their goals seemingly compete with each other, the maximization underlying the optimal classification model and the JCM are comparable. That is, assigning an applicant to the job with the highest predicted performance or criterion estimate corresponds to an individual choosing an alternative with maximum utility. In the next section, we will transform the criterion estimate to obtain the same mathematical structure as the utility equations, then solve the classification problem using DCM estimation procedures.

DCM Approximation of Brogden Classification Analysis

In this section we describe how the traditional Brogden classification problem can be formulated using a DCM framework. We will begin by summarizing De Corte's implementation of Brogden's augmented criterion estimate assignment strategy and then propose an alternative approximate implementation that borrows ideas from DCM methodology. In a later section, we will draw insights from this approximate implementation to propose an approach for conducting classification analysis that is more robust with regard to distributional assumptions and can accommodate important policy constraints and applicant preferences.

Multivariate Normal Implementation of Brogden Optimal Classification

The following discussion summarizes De Corte's implementation of Brogden's optimal classification strategy. Our development uses a slightly different approach but closely follows the main ideas in his work. With respect to notation used throughout this report, we will use upper case letters to represent random variables and lower case letters to represent observed (sample) values; random vectors and random values will be presented in bold font.

Classification Outcome Vector

Let the random variable X_{ij} denote the estimated or predicted criterion score of the *i*th applicant for the *j*th job. Also assume that the vector of estimated criterion scores for a randomly selected applicant, $\boldsymbol{X}_i = \langle X_{i1}, \ldots, X_{im} \rangle$, is distributed as multivariate normal. De Corte developed an approach for the general optimal classification problem based on Brogden's augmented criterion estimate assignment algorithm. This algorithm modifies the *j*th criterion estimate for classification purposes by adding a job-specific constant A_j to obtain the augmented criterion estimate $X_{(a)ij} = A_j + X_{ij}$. The algorithm assigns the *i*th applicant to the *k*th job if $X_{(a)ik} > X_{(a)ij}$ for all other jobs $j \neq k$ (i.e., $X_{(a)ik}$ is the maximum modified assignment criterion). Brogden showed that given suitably chosen job specific constants (A_js), applying the algorithm to a large pool of applicants will approximately produce the desired job allocation quota.

Brogden's optimal classification assignment strategy can be viewed as mapping from the *m*-dimensional random vector of criterion estimates, $\langle X_{i1}, \ldots, X_{im} \rangle$, to the 2-dimensional random vector $\langle X_i^{(h)}, K_i^{(h)} \rangle$, where $X_i^{(h)}$ is a random variable that takes its value from the criterion estimate for the optimal job (i.e., job with the highest augmented criterion estimate) of the *i*th applicant, and $K_i^{(h)}$ is a random variable whose value is the index of the optimal job. Note that $\langle X_i^{(h)}, K_i^{(h)} \rangle$ completely describes the outcome of the algorithm for the *i*th applicant. The joint probability density function (PDF) of this outcome vector can be expressed as the product of the probability of observing the criterion estimate x for the *k*th job and the conditional probability that this observed score is the highest among all *m* augmented criterion estimates. That is,

$$f\left(K_i^{(h)} = k, X_i^{(h)} = x\right) = P\left(K_i^{(h)} = k, X_{ik} = x\right)$$
$$= f_{X_k}(x)P\left(K_i^{(h)} = k \middle| X_{ik} = x\right)$$
$$= f_{X_k}(x)P\left(X_{ij} + A_j < x + A_k \middle| X_{ik} = x, k \neq j\right)$$

The two probability factors in the last line above were completely specified by De Corte. The first factor is simply the univariate normal PDF corresponding to the *k*th criterion estimate evaluated at x. To evaluate the second factor, De Corte used the lower tail probability of the $(m-1)$-dimensional normal distribution obtained by conditioning the *k*th criterion estimate to x.[1] For discussion purposes we will leave the joint probability in the above final form.

Job Quota Constraints

We now obtain two key components in De Corte's approach. The first component is the system of nonlinear equations representing job quota constraints. These equations equate the job quotas to the percentages of applicants that the augmented criterion estimate algorithm assigns to the *m* jobs. These percentages are just the marginal probabilities of $K_i^{(h)}$ that can be obtained by integrating the joint PDF of the outcome vector over all possible values of $X_i^{(h)}$. Thus, the *m* job quota equation constraints are

$$\int_{-\infty}^{+\infty} f_{X_{ik}}(x)P\left(X_{ij} + A_j < x + A_k \middle| X_{ik} = x, k \neq j\right)dx = q_k, k = 1, \ldots, m \quad (1)$$

where the left-hand side integral equals the marginal PDF $f_{K_i^{(h)}}(k)$. Note that the variables in the above system of nonlinear equations are the unknown job-specific constants.

Calculating MPP

A second component in De Corte's approach is the marginal PDF of the predicted criterion scores of applicants on their optimal jobs (i.e., the $X_i^{(h)}$s). We need this PDF to evaluate the MPP. It can be derived by summing the joint PDF of the outcome vector $\langle X_i^{(h)}, K_i^{(h)} \rangle$ over all possible values of $K_i^{(h)}$. Conceptually, this means anyone of the *m* jobs can be the optimal job. Thus the marginal PDF of $X_i^{(h)}$ evaluated at x is

[1] Note that in De Corte's derivation x is the observed value of the augmented criterion score, while in our presentation above x is the observed value of the unmodified criterion score.

$$f_{X_i^{(h)}}(x) = \sum_{k=1}^{m} f\left(K_i^{(h)} = k, X_i^{(h)} = x\right)$$
$$= \sum_{k=1}^{m} f_{X_k}(x) P\left(X_{ij} + A_j < x + A_k \middle| X_{ik} = x, k \neq j\right)$$

Note that this marginal PDF involves job specific constants $A_j$as unknown parameters. To completely define this PDF, we solve the job quota constraint system of nonlinear equations for the A_js.

Having defined the marginal PDF of the predicted criterion scores of applicants on their optimal jobs, the MPP may now be evaluated by taking the expectation of $X_i^{(h)}$

$$MPP = E\left(X_i^{(h)}\right) = \sum_{k=1}^{m} \int_{-\infty}^{+\infty} x f_{X_k}(x) P\left(X_{ij} + A_j < x + A_k \middle| X_{ik} = x, k \neq j\right) dx \qquad (2)$$

It is also informative to express the MPP as the weighted average of job − level MPPs

$$MPP = \int_{-\infty}^{+\infty} x \sum_{k=1}^{m} f\left(K_i^{(h)} = k, X_i^{(h)} = x\right) dx$$
$$= \sum_{k=1}^{m} q_k \int_{-\infty}^{+\infty} x \frac{f\left(K_i^{(h)} = k, X_i^{(h)} = x\right)}{q_k} dx$$
$$= \sum_{k=1}^{m} q_k \int_{-\infty}^{+\infty} x f\left(X_i^{(h)} = x \middle| K_i^{(h)} = k\right) dx$$
$$= \sum_{k=1}^{m} q_k E\left(X_i^{(h)} \middle| K_i^{(h)} = k\right)$$

We obtained the third line by noting that $f_{K_i^{(h)}}(k) = q_k$ and applying the definition of a conditional PDF. The conditional average performance or conditional MPP, $E\left(X_i^{(h)} \middle| K_i^{(h)} = k\right)$, is useful for assessing classification efficiency pattern across jobs, while the conditional PDF $f\left(X_i^{(h)} = x \middle| K_i^{(h)} = k\right)$ is useful for describing the distribution of optimal predicted performance scores of applicants in the *k*th job.

MMNL Approximation of Brogden Classification Framework

We next propose an alternative approach to the general optimal classification problem by borrowing ideas from DCM methodology. This alternative approach is presented below as an approximate implementation of Brogden's optimal classification framework. In the next section, we will draw insights from the analytic results derived below to propose a more robust and flexible approach for conducting classification efficiency analysis.

Modified Augmented Criterion Estimate

We begin by proposing the modified augmented criterion estimate for the *j*th job of the *i*th applicant given by

$$X_{(a')ij} = \lambda\left(A_j + X_{ij}\right) + E_{ij}$$

where the E_{ij}s are independent standard Gumbel random variables, λ is an arbitrary scaling parameter, and (as before) X_{ij}s are the criterion estimates and A_js are job-specific constants. Note that $X_{(a')ij}$ is simply a rescaled version of the original augmented criterion estimate, $X_{(a)ij}$, with an additional Gumbel random variable term. In general the rank ordering of $X_{(a')ij}$s can differ from that of $X_{(a)ij}$s and, therefore, can lead to a different optimal job for the *i*th applicant. However, we can make λ suitably large so that $\lambda\left(A_j + X_{ij}\right) = \lambda X_{(a)ij}$ predominantly determines the rank ordering of $X_{(a')ij}$s. Since scaling does not alter the rank ordering of jobs, a modified optimal classification algorithm based on $X_{(a')ij}$s can approximate Brogden's optimal classification algorithm to a desired degree of accuracy (i.e., both will produce the same applicant-job assignments) by choosing an arbitrarily large value of λ. The idea behind this modified augmented criterion estimate approximation is based on McFadden and Train (2000).

We make the following observation before deriving job quota constraint equations and the PDF of the optimal criterion estimate of applicants under the modified classification algorithm. Given fixed criterion estimates, $\boldsymbol{X}_i = \boldsymbol{x}_i$, the form of the modified augmented criterion estimates $X_{(a')ij}$s follow that of the utility equations of a DCM with an "observed" linear component, represented by $\lambda\left(A_j + X_{ij}\right)$, and an "unobserved" component, represented by the Gumbel random variable E_{ij}. This means that the conditional probability of a specific job having the highest modified augmented criterion estimate given $\boldsymbol{X}_i = \boldsymbol{x}_i$ has the form of the multinomial logit (MNL) probability

$$P\left(X_i^h = x_k \middle| \boldsymbol{X}_i = \boldsymbol{x}_i\right) = \frac{\exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)}$$

Job Quota Constraints

Under the modified optimal classification algorithm, it can be shown (see Appendix A) that the job quota constraints can be represented by the system of nonlinear equations below, with variables A_js:

$$\int_{x_i \in R^m} \frac{\exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_{X_i}(\boldsymbol{x}_i) d\boldsymbol{x}_i = q_k \,, k = 1, \ldots, m \tag{3}$$

The system of equations (3) corresponds to (1) in the MVN model. Note that the left-hand side of each equation above follows the form of a MMNL probability model, with utility equations composed of alternative-specific constants (ASC) λA_js, and normally distributed "error components" x_{ik}s. The nonlinear equations above comprise the first-order condition for the maximum likelihood estimation (i.e., first derivative of the log-likelihood function equated to zero) for such an MMNL model. In the job choice modeling context, the left-hand-side is the predicted share of the *k*th job in the population, which is equated to q_k, the percentage of applicants that actually chose the *k*th job. These observations imply that we can solve for the job specific constants that satisfy the job quota constraints using an MMNL parameter estimation algorithm. Biogeme (Bierle, 2003) model files for carrying out the estimation of selected optimal classification problems are specified in Appendix B.

While based on a somewhat awkward interpretation of criterion estimates (x_{ik}s) as "error components," the above observation provides a connection between Brogden's optimal job classification and DCM from a computational viewpoint. In the next section, we provide a more useful conceptual relationship between optimal classification and DCM that will lead to a more practical framework for optimal classification analysis.

Calculating MPP

As in de Corte's method, the MPP can be evaluated by taking expectations using the PDF of the estimated criterion score corresponding to the optimal job of a randomly chosen applicant under the modified classification algorithm. The PDF under the modified classification algorithm is given by

$$f^*_{X_i^{\mathrm{th}}}(x) = \sum_{k=1}^{m} f_{X_k}(x) \int_{x_i \in R^{m-1}} \frac{\exp\left(\lambda\left(A_k + x\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_{\boldsymbol{X}_{(k)}|X_k}\left(\boldsymbol{x}_{(k)}|x\right) d\boldsymbol{x}_{(k)}$$

Evaluating expectations using this PDF, we obtain the following expression for the MPP (see Appendix A):

$$MPP = \sum_{k=1}^{m} \int_{x_i \in R^m} x \frac{\exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x} \tag{4}$$

Note that MPP (4) corresponds to (2) under the MVN model. The PDF $f^*_{X_i^h}(x)$ involves the unknown job specific constants as parameters. This PDF is completely specified by solving the job quota system of nonlinear equations for the unknown constants.

Comparison Between MMNL and MVN Optimal Criterion PDFs

We briefly comment on the relationship between the PDFs of the optimal criterion estimate derived by De Corte under the original Brogden classification algorithm and the corresponding PDF under the modified classification algorithm. Both PDFs $f^*_{X_i^h}(x)$ and $f_{X_i^h}(x)$ have the same general form but the terms inside their summation differ slightly. The integration in the final expression for $f^*_{X_i^h}(x)$ is the conditional MMNL probability of identifying the *k*th job as optimal with a normal mixing distribution of dimension (*m*-1), given by $f_{\boldsymbol{X}_{(k)}|X_k}(\boldsymbol{x}_{(k)}|x)$. It corresponds to the lower tail probability in $f_{X_i^{(h)}}(x)$, which equals the probability of identifying the *k*th job as optimal using the conditional probit model $f_{\boldsymbol{X}_{(k)}|X_k}(\boldsymbol{x}_{(k)}|x)$. Apart from missing the MNL probability, the integral in $f_{X_i^h}(x)$ is evaluated only on the subspace $\{x_j + A_j < x + A_k \,; j \neq k\}$. This is because under Brogden's original algorithm, the remaining (*m*-1) criterion estimates are restricted once the optimal job is identified, while under the modified algorithm the (*m*-1) criterion estimates can be any point in R^{m-1} since the Gumbel random variable components can ultimately determine the optimal job. Lastly, it is easy to verify that

$$\lim_{\lambda\to\infty} \frac{\exp\left(\lambda(A_k + x)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} = 1_{(x_j + A_j < x + A_k \,; j \neq k)}$$

so that as λ becomes large the integration in $f^*_{X_i^h}(x)$ becomes approximately equal to $f_{X_i^{(h)}}(x)$, as expected.

Simultaneous Selection-Classification Problem

For completeness, we derive the MMNL approximation model for evaluating MPP under simultaneous optimal selection-classification using our proposed algorithm. In selection-classification, the sum of the *m* job quotas is less than 100 percent, with the remainder equal to the rejection percentage. In his multivariate normal implementation, De Corte handled the selection-classification problem by introducing a cutoff value for the marginal PDF of X_i^h. This cutoff value became an additional variable in the system of nonlinear equations representing job quota constraints (see equation (9) in De Corte), along with the *m* job specific constants. For the modified classification algorithm, we expand the RUM interpretation of the optimal assignment rule. We accomplish this by including an "auxiliary job" where non-accessions or rejected applicants are "assigned." Using index $j = 0$ to identify the non-accession job, the augmented criterion estimate (or utility) corresponding to this non-accession job is given by

$$X_{(\alpha')i0} = \lambda A_0 + E_{i0}$$

where A_0 is an additional job specific constant and E_{i0} is an extra standard Gumbel random variable that is independent of the other *m* E_{ij}s. Behaviorally, to identify the optimal job for an applicant, we simply treat the non-accession just like any of the other *m* jobs. In this case, the non-accession job is the "optimal job" for an applicant if

$$\lambda(A_j + X_{ij}) + E_{ij} < \lambda A_0 + E_{i0}\ ;\ j \neq 0$$

Likewise, we assign the applicant to the *k*th job if

$$\lambda A_0 + E_{i0} < \lambda(A_k + X_{ik}) + E_{ij}$$

and

$$\lambda(A_j + X_{ij}) + E_{ij} < \lambda(A_k + X_{ik}) + E_{ij}\ ; j \neq k$$

Using the above optimal assignment rules, the job quota constraint for the selection-classification problem becomes

$$\int_{x_i \in R^m} \frac{\exp(\lambda A_0)}{\exp(\lambda A_0) + \sum_{j=1}^{m} \exp\left(\lambda(A_j + x_{ij})\right)} f_{X_i}(\boldsymbol{x_i}) d\boldsymbol{x_i} = q_0$$

$$\int_{x_i \in R^m} \frac{\exp\left(\lambda(A_k + x_{ik})\right)}{\exp(\lambda A_0) + \sum_{j=1}^{m} \exp\left(\lambda(A_j + x_{ij})\right)} f_{X_i}(\boldsymbol{x_i}) d\boldsymbol{x_i} = q_k\ ;\ k = 1, \ldots, m$$

where $q_0 = 1 - \sum_{j=1}^{m} q_j$. Note that the left-hand side expressions are just the probabilities of assigning an applicant to the non-accession job and the *k*th job. As in the classification-only problem, solving for the $(m + 1)$ job-specific constants above can be viewed as an MMNL model parameter estimation problem. Mathematically, the constant q_0 is equivalent to the cut off value in De Corte's approach. It might be worth investigating if other RUM models can be applied to solve other variations of classification (e.g., multi-stage, etc.)

Example utility equations and input data requirements of an MMNL model for approximating Brogden optimal classification using the Biogeme model file syntax and data format are described in Appendix B.

Empirical (Sample Based) Optimal Classification Analysis

In this section we draw insights from analytical results derived in the previous section to propose a more practical approach for conducting classification efficiency analysis. We will show that this approach can readily accommodate personnel classification policy constraints and applicant preferences, and provide a more accurate representation of the "true" distribution of the applicant pool. As such, it can produce more realistic results compared to the "exact" MVN and MMNL methods presented in the previous section.

Limitations of a MVN Population Classification Model

Brogden's optimal classification method assumes that the sample size approaches infinity and that job criterion estimates are distributed as multivariate normal. In applications, MPP calculations are based on a multivariate normal population with mean and covariance computed from a large sample of criterion estimates. Both De Corte's MVN implementation and our proposed MMNL approximation replaced the sample of criterion estimates from actual applicants with a multivariate normal population. In other words, both implementations use a smooth distribution to represent the true applicant distribution.

Assuming a multivariate normal distribution for the criterion estimates made classification analysis mathematically tractable. There are two key problems with this approach, however. First, the true distribution of criterion estimates is usually not exactly normal. Because of the nature of optimization, the analysis can be sensitive to departures from normality, especially as the number of jobs increases. In Army applications with a very large applicant sample, discarding the sample data and instead using a multivariate normal distribution to *approximate* the distribution of applicant criterion estimates is wasteful, when the sample of criterion estimates computed from the data offers a more accurate representation of the true distribution.[2] Classification analysis that directly uses the sample of criterion estimates will be more robust to departures from normality. Second, it is difficult to mathematically include real world classification constraints, such as cut scores and gender restrictions, under the multivariate normal assumption. Classification analysis that can accommodate such constraints will give more realistic policy guidance.

MNL Model for Optimal Classification

We now obtain an approach for classification efficiency analysis that takes advantage of sample data (i.e., not based on an assumed mathematical distribution). We first derive a model that relaxes the multivariate normal distribution assumption for applicant criterion estimates without eligibility constraints.

Implied MNL Model

An empirical or sample data approach for optimal classification analysis is implied in the MMNL approximation model presented in the previous section. To derive this empirical

[2] Classification efficiency is concerned with criterion estimates and not with actual criterion values.

approach we reexamine the system of nonlinear equations (3) used to represent job quota constraints under the MMNL approximation model. These equations are reproduced below:

$$\int_{x_i \in R^m} \frac{\exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_{X_i}(\boldsymbol{x}_i) d\boldsymbol{x}_i = q_k, k = 1, \ldots, m$$

As mentioned earlier, these equations are equivalent to the first-order conditions for the Maximum Likelihood Estimator (MLE) of an ASC-only MMNL model with correlated "error components" (i.e., x_{ij}s) that are distributed as multivariate normal and observed job choice percentages given by q_ks. We used this interpretation to derive the MMNL maximum likelihood approach for computing the job specific constants in the modified augmented criterion estimate. In the MMNL classification model, the population of applicant criterion estimates was represented by the multivariate normal distributed error components. To satisfy the job quota constraints, a synthetic estimation dataset was constructed with number of observations equal to number of jobs. Each observation in this dataset represents a unique chosen job, with observation weight equal to the corresponding job quota.

When a large sample of applicants is available, we can replace the multivariate normal distributed error components in the left-hand-side (LHS) of the job quota constraint equations with the applicants' criterion estimates. The original MMNL approximating model now becomes a simple MNL model with only job-specific constants. We construct the needed synthetic estimation dataset that satisfies the job quota constraints by: (a) replicating each applicant observation in the sample as many times as the number of jobs, with each replicate corresponding to the applicant choosing a unique job; and (b) setting the weight for each observation replicate to be equal to the quota for the job chosen in the replicate. The second step ensures that the job proportions of job "choices" in the synthetic estimation data correspond to the target job quotas.

Replacing integration with respect to the MVN distribution with a summation over the sample observation replicates, the job quota constraint equations becomes

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} = q_k, k = 1, \ldots, m \qquad (5)$$

where n is the total sample size.[3] Note that for each k the left-hand side is simply the average of the optimal assignment probabilities of applicants in the sample for the *k*th job. The preceding equations are the first-order conditions for an empirical MNL model with x_{ij}s as alternative-specific predictors with fixed coefficient ($\beta = 1$), A_js as unknown ASCs, and q_ks as observed

[3] For given job *k*, the LHS involves the sum of weighted probabilities for the *k*th job across $n \times m$ observation replicates. However, since the weights for the *m* replicates of an applicant add up to one, the sum of weighted probabilities across observation replicates simplifies to the sum of probabilities across unreplicated observations as shown above.

job choice percentages. We can solve for the A_js by applying the maximum likelihood estimation for this MNL model using the aforementioned synthetic estimation data.

To obtain the computational formula for the MPP under the empirical MNL approximation model, we examine the expression of the MPP under the MMNL population based approximation model:

$$MPP = \sum_{k=1}^{m} \int_{x_i \in R^m} x \frac{\exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_X(\boldsymbol{x})\, d\boldsymbol{x}$$

Again, replacing the integration with respect to the multivariate normal PDF $f_{X_i}(\boldsymbol{x_i})$ with summation over the sample data, we obtain the computational expression below for the MPP:

$$MPP = n^{-1} \sum_{k=1}^{m} \sum_{i=1}^{n} x_{ik} \frac{\exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} \tag{6}$$

In other words, the MPP in the MNL model is simply the weighted average of criterion estimates across all applicants and jobs, with MNL optimal job assignment probabilities as weights. Note that the general form of this formula is the same as the formula for computing the average bonus in JCM models (Diaz, Ingerick, & Sticha, 2007b; Diaz, Sticha, Hogan, Mackin, & Greenston, 2012). The difference is that here probabilities are derived from optimal assignment rules while the JCM probabilities were derived from actual applicant choices.

Illustration Using Simulated Data

We illustrate the MNL model for classification analysis using criterion estimates that were simulated from a multivariate normal distribution with zero mean and covariance matrix

$$\mathbf{V} = \begin{bmatrix} 0.5 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.6 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.7 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.5 \end{bmatrix}$$

The Biogeme MNL model file and input data specification are described in Example 4 of Appendix B. Table 1 shows the job specific constants and job level and overall MPP obtained under the sample based MNL model for optimal classification analysis. For comparison, we also show the results obtained for the population based MVN and MMNL model for the correlated criterion estimates examples in Appendix B. The constants and MPP shown in Table 1 for the MVN population model were computed using implementation of De Corte's method. The constants for the MMNL population model were computed using Biogeme with model and data files as specified in Example 2 of Appendix B. (We did not implement the procedure for computing the MPP under the MMNL approximation model. The MPPs under the MMNL model should be very close to those computed using the MNL sample based model.) The constants and MPP for the empirical MNL model were computed as described in this section.

Overall, calculated values for the job specific constants closely match across the three methods. The job level and overall MPPs calculated under the MVN population model and MNL sample based model also match very closely. These observations are not surprising since the sample data used in the MNL model were simulated from the same multivariate normal population assumed in MVN example. In the next examples we will compare the results of the empirical MNL model to results of the population based MVN and MMNL models when accounting for eligibility requirements.

Table 1.*Comparison of Sample Based MNL Results with Population Models*

Model	Job1	Job2	Job3	Job4	All
Job Specific Constants					
MNL	0.2719	0.1089	-0.1273	-0.2758	
MVN	0.2719	0.1204	-0.1298	-0.2625	
MMNL	0.2719	0.0966	-0.1300	-0.2730	
Mean Predicted Performance					
MNL	0.3947	0.5972	0.8438	0.7385	0.5796
MVN	0.3996	0.5940	0.8484	0.7489	0.5827
MMNL					

Including Applicant Eligibility Constraints

Current applications of Brogden classification efficiency analysis ignore important personnel classification constraints, such as gender restrictions and cut scores. This limitation is inherent in a method that is mainly based on an assumed multivariate normal distribution. It can substantially impact the results of the analysis, depending on how the true distribution of criterion estimates differs from a multivariate normal distribution. For example, gender restriction makes combat jobs off limit to a sizeable subset of the applicant population, effectively reducing size of the optimization space. Therefore, ignoring this restriction will likely produce higher overall MPP than possible operationally. In practice, this could mean arriving at inaccurate or incorrect conclusions regarding the incremental classification efficiency potential of experimental predictor batteries for the full accession cohort or for specific jobs. In this section, we will generalize the sample data based job quota and MPP equations (5) and (6) to the case where applicant eligibility constraints are included in the problem.

Modified MNL Model

Limitations of the Brogden optimal classification approach related to applicant eligibility can be easily fixed using the empirical MNL optimal classification model. In an MNL model, as with all DCMs, the analysis data can include indicator variables identifying which alternatives were available to individual decision makers. Using these indicator variables, we can readily

include policy constraints related to applicant job eligibility in the MNL classification model. Making some jobs unavailable to some applicants requires adjustments to the weights of observation replicates to ensure that the proportions of "chosen" jobs are equal to the targeted job quotas. It also requires modification to the job quota constraint equations because only those applicants who are eligible for a given job contribute to the quota for the job.

The adjusted replicate weights are constructed in two steps as follows. First, we obtain provisional replicate weights by proportionately allocating the contribution of an applicant to job quotas based on his eligibility profile. Using the indicator variable δ_{ir} to denote the *i*th applicant's eligibility for the *r*th job, these provisional replicate weights are given by

$$q_{ir}^{*} = \frac{\delta_{ir} q_r}{\sum_{j=1}^{m} \delta_{ir} q_r}$$

Note that $q_{ir}^{*} = 0$ if the applicant is not eligible for the *r*th job. Second, we adjust the provisional weights so that the sum of the final replicate observation weights for the *j*th job is proportional to the quota for the job, yielding final replicate observation weights given by

$$w_{ir} = q_{ir}^{*} \frac{q_r}{\sum_{i=1}^{n} q_{ir}^{*}}$$

Note that $\sum_{i=1}^{n} w_{ir} = q_r$, $r = 1, \ldots, m$ as desired.

We can now construct the job quota constraint equations (i.e., first-order MLE conditions) for the MNL optimal classification model with applicant job eligibility constraints by summing probabilities across the $n \times m$ replicate observations, as follows:

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{r=1}^{m} w_{ir} \left[\frac{\delta_{ik} \exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \delta_{ij} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} \right] = q_k, k = 1, \ldots, m \tag{7}$$

The overall MPP is calculated under the modified MNL model with eligibility constraints by simply taking the weighted average of criterion estimates across replicate observations, with weights equal to the products of the optimal job assignment probabilities and replicate weights. This weighted average is of the form

$$MPP = m^{-1} \sum_{k=1}^{m} \sum_{i=1}^{n} \sum_{r=1}^{m} x_{ik} \left[w_{ir} \frac{\delta_{ik} \exp\left(\lambda\left(A_k + x_{ik}\right)\right)}{\sum_{j=1}^{m} \delta_{ij} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} \right] \tag{8}$$

The summation above is simply the sum of the product of applicant criterion estimates and their corresponding assignment probabilities, using only applicants eligible for each job. Note that the equations (7) and (8) simplify to equations (5) and (6) if each applicant is eligible for all *m* jobs.

Examples Using Actual Data

We present illustrative examples of the sample based MNL optimal classification model using data from the Tier One Performance Screen (TOPS) initial operational test and evaluation study (Knapp & Heffner 2012). In these examples we evaluate the classification potential of ASVAB and TAPAS predictors for minimizing 6-month attrition in four MOS, under pure classification and selection-classification systems, while accounting for classification policy requirements. The two classification requirements considered are MOS minimum aptitude area (AA) score and gender restriction. Table 2 shows the four MOS, their minimum AA score requirement, and gender restriction.

The initial steps in applying the MNL optimal classification model in this problem are as follows. The first step is to estimate the probability of 6-month attrition, the criterion estimate of interest, using data from each of the four MOS. To evaluate incremental classification efficiency of TAPAS beyond the ASVAB, we estimated two logistic models for each MOS, one using only ASVAB scores as predictors and the other using both ASVAB and TAPAS scores as predictors. The second step is to apply the estimated logistic models to a larger sample of applicants to obtain an empirical distribution of criterion estimates in the applicant population. We used a large sample of accessions (n=62,155) from the TOPS study to represent the applicant input. Table 2 shows the sizes of MOS samples used to develop the 6-month attrition logistic probability models, observed attrition in the sample, and pseudo R2 of the estimated models under the two sets of predictors corrected for restriction in range relative to the full sample.[4] Overall observed attrition for all four MOS combined is 11.5%.

Note that with attrition probabilities as criterion estimates, the normal distribution assumption in de Corte's approach no longer holds. To get around this problem, Trippe, Diaz and Ingerick (2012) used the linear predictor function in the logistic model to determine optimal classification of applicants. They then evaluated the mean attrition for each MOS by taking expectation with respect to the distribution of linear predictors of optimally classified applicants in each MOS. This is a reasonable approach because the linear predictor is monotonically related to the probability of attrition. Because the linear predictor is a weighted sum of continuous variables (i.e., ASVAB and TAPAS scores) the normal distribution assumption is also tenable.

Under the MNL optimal classification model the distribution of the criterion estimates is determined empirically from sample data. Therefore it is possible to directly use attrition probability estimates to determine the optimal classification of applicants. In the first set of examples below, applicants are classified using the linear predictor under four classification conditions based on rejection rate and classification policy constraints. These conditions are shown in Table 3 in order of increasing number of constraints. In the second set of examples applicants are classified directly using attrition probability estimates.

[4] Correction for restriction in range was carried out by applying McKelvey and Zavoina (1975) pseudo R2 to the full sample, $R^2 = \frac{var(BX)}{var(BX)+\pi^2/3}$, where BX is the linear predictor of the logistic model $P(y=1) = \frac{exp(BX)}{1+exp(BX)}$.

Table 2. ***Eligibility Requirements and 6-Month Attrition for Selected MOS***

	Eligibility Conditions		6-Month Attrition			
					Model Pseudo R2	
MOS	Cut Score	Gender	N	Actual Attrition	ASVAB	ASVAB+TAPAS
11B	CO>=87	Male Only	4702	12.14	0.0257	0.0710
31B	ST>=91	NA	264	12.12	0.0851	0.2130
68W	ST>=101	NA	990	8.08	0.0470	0.1203
88M	OF>=85	NA	662	11.33	0.0466	0.1572

Table 3. ***Classification Conditions***

Reject	Eligibility	Description
0%	None	Pure classification, ignoring cut scores and gender restriction
10%	None	Selection-Classification, ignoring cut scores and gender restriction
10%	CS	Selection-Classification, enforcing MOS cut scores only
10%	CS + (11B Male)	Selection-Classification, enforcing MOS cut scores and gender restriction

Table 4 shows the mean 6-month average attrition estimate for the entire sample and by MOS when applicants are optimally classified using the linear predictor under different conditions and predictors, with allocation percentages of 46.2%, 18.9%, 22.9%, and 13.0%, for 11B, 31B, 68W, and 88M, respectively. The table is organized to facilitate comparison of results between classification conditions for a given set of predictors. Compared to actual attrition percentages in the sample, optimally classifying applicants using only ASVAB scores with no policy constraints produced substantial improvements in mean attrition probabilities for the entire sample and by MOS except for 11B. This result is not surprising given that 11B accounts for close to half of the sample and has a pseudo-R2 that is practically zero.

Rejecting 10% of applicants and optimally classifying the remaining 90% with no constraints reduced overall and MOS attrition probabilities by about half percentage point.[5] Adding a cut score constraint when classifying 90% of the applicants increased overall attrition from 8.0% to 8.1%. While negligible this is the anticipated direction in mean attrition due to addition of the cut score constraint. Note that mean attrition for 68W increase from 6.0% to 6.1% while mean attrition for 31B increase from 5.6% to 5.8%. Again, while negligible, changes observed for 31B and 68W are meaningful given that both MOS use the same AA composite (ST), with 68W having the higher cut score (ST>=101) compared to 31B (ST>=91). In other words, 31B and 68W more directly compete with each other than with 11B and 88M, with 68W having priority to better applicants (with ST >=101) compared to 31B. Including a gender restriction in addition to the cut score constraint produced an additional increase in overall attrition, from 8.1% to 8.2%, with 11B having the largest increase (10.5% to 10.8%) as expected.

[5] A small rejection rate of 10% was used because 3% of accessions do not qualify for any of the four MOS when using both cut score and gender eligibility constraints.

Table 4. ***Summary of Classification Potential of the ASVAB and TAPAS for Minimizing 6-Month Attrition Using Logistic Linear Composite Optimal Classification by Condition***

Reject	Eligibility	11B	31B	68W	88M	ALL
ASVAB Only						
0%	None	11.1	6.0	6.5	6.6	8.5
10%	None	10.5	5.6	6.1	6.2	8.0
10%	CS	10.5	5.8	6.0	6.2	8.1
10%	CS + (11B Male)	10.8	5.8	6.1	6.3	8.2
ASVAB+TAPAS						
0%	None	12.1	3.6	5.3	3.3	7.9
10%	None	10.7	3.2	4.7	2.9	7.0
10%	CS	10.9	3.3	4.7	3.0	7.1
10%	CS + (11B Male)	11.1	3.4	4.8	3.1	7.3
Incremental Efficiency of TAPAS Over ASVAB						
0%	None	-1.0	2.4	1.2	3.3	0.7
10%	None	-0.2	2.5	1.4	3.2	1.1
10%	CS	-0.4	2.5	1.3	3.3	1.0
10%	CS + (11B Male)	-0.4	2.5	1.3	3.2	1.0

The second set of results in Table 4 summarizes the potential classification benefits of using both TAPAS and ASVAB scores on 6-month attrition. Optimally classifying applicants using TAPAS and ASVAB scores yielded substantial reduction in average attrition over the classification model based only on ASVAB for the entire sample and by MOS with exception of 11B. The reduction in estimated attrition rate is substantial (about 50%) for MOS 31B and 88M across all conditions. MOS 68W showed a modest decrease in estimated attrition. The estimated attrition for MOS 11B unexpectedly went up by less than one percentage point across all conditions. The general pattern of small increases in attrition rates as cut score and gender restriction constraints are included in the classification model is the same as those observed in the model using only ASVAB.

Table 5 shows the mean 6-month attrition probabilities for the entire sample and by MOS when applicants are optimally classified directly using attrition probability estimates based on the logistic model. The first set of results in Table 5 shows the potential classification benefits of using only ASVAB scores as predictors under different conditions. Using the logistic probability estimates of attrition to optimally classify applicants produced approximately the same estimated attrition rates for the overall sample across conditions, compared to classification based on the linear predictor. The estimated attrition rates by MOS when classifying using attrition probability estimates are also generally comparable to the attrition rates when classification is based on the linear predictor (see Table 4). However, the attrition rates for MOS 31B, 68W, and 88M increase by about a half percentage point, while the attrition rates for MOS 11B decrease by 0.6 to 0.9 percentage points.

Table 5. ***Summary of Classification Potential of the ASVAB and TAPAS for Minimizing 6 Month Attrition Using Logistic Probability Optimal Classification by Condition***

Reject	Eligibility	11B	31B	68W	88M	ALL
ASVAB Only						
0%	None	10.2	6.5	7.3	6.9	8.4
10%	None	9.9	6.0	6.7	6.4	8.0
10%	CS	9.9	6.1	6.5	6.6	8.0
10%	CS + (11B Male)	10.1	6.2	6.7	6.7	8.2
ASVAB+TAPAS						
0%	None	9.7	4.9	6.4	5.1	7.5
10%	None	9.0	4.1	5.5	4.1	6.7
10%	CS	9.1	4.3	5.4	4.4	6.8
10%	CS + (11B Male)	9.2	4.4	5.6	4.5	6.9
Incremental Efficiency of TAPAS Over ASVAB						
0%	None	0.5	1.6	0.9	1.8	1.0
10%	None	0.9	1.9	1.2	2.4	1.3
10%	CS	0.8	1.9	1.2	2.2	1.2
10%	CS + (11B Male)	0.9	1.8	1.1	2.2	1.3

The second set of results in Table 5 summarizes the potential classification benefits of using both TAPAS and ASVAB scores on 6-month attrition when applicants are classified directly using attrition probability estimates. Compared to classifying applicants using the linear predictor, the overall attrition estimates are lower by about 0.4 percentage points when classifying applicants directly using the estimated attrition probabilities. At the MOS level, attrition rates for MOS 31B, 68W, and 88M increased, approximately 1 to 2 percentage points, while the attrition rates for MOS 11B decreased by 1.7 to 2.4 percentage points, when classifying applicants using attrition probability estimates compared to classifying using linear predictor.

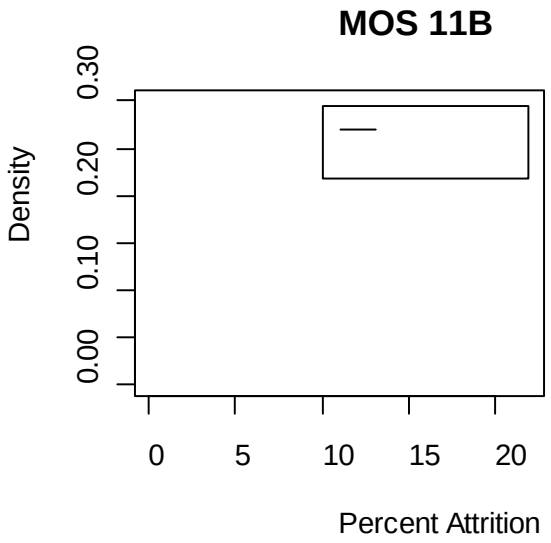
Comparing the results for the two sets of predictors in Table 5 shows that optimally classifying applicants using TAPAS and ASVAB scores yielded substantial reduction in average 6-month attrition for the entire sample and all MOS over the classification model based only on ASVAB. Note that unlike classification using the linear predictor, classifying applicants using the estimated attrition probability shows that MOS 11B also benefits from the inclusion of TAPAS in addition to the ASVAB, along with a higher incremental improvement for the entire sample. In other words, classifying applicants directly using attrition probability estimates produced stronger evidence of classification efficiency compared to classification using the linear predictor.

Using a linear predictor appeared to have favored MOS with lower estimated attrition probabilities (or higher probabilities of completing the first term) at the expense of MOS 11B. This “bias” is likely due to the fact that a linear predictor is no longer linearly proportional to logistic probability estimates for completing the term in the upper tail of the distribution. For the

same differences in linear predictors in the tails and middle portion of the logistic model, the corresponding differences in probabilities between jobs are small. Therefore, classifying applicants using a linear predictor would tend to prioritize better applicants to jobs with lower attrition estimates even if they can contribute better to lowering overall attrition in jobs with higher attrition. Figure 1 graphically shows the bias of classification using a linear predictor (solid lines) over direct classification using attrition probability estimates (dashed lines). Direct classification shifted the distribution of attrition probability estimates for MOS 31B, 68W, and 88M to the right, with a corresponding shift to the left in the distribution for MOS 11B. The net effect of these shifts is higher overall classification efficiency, as noted in observations from Tables 4 and 5.

In sum, we applied the MNL optimal classification model to evaluate the potential classification efficiency of ASVAB and TAPAS when classification policy constraints are included in the model. While the differences in average attrition estimates as policy constraints were included in the classification model were negligible, the pattern of changes were consistent with anticipated effects of the cut score and gender restriction on the MOS in the example. We expect these effects to be stronger with more valid criterion models and greater differentiation in classification policy. For example, cut score policy differentiation between MOS 31B (ST>=91) and 35T (ST>=112) is greater compared to that between MOS 31B and 68W (ST>=91). We also demonstrated that, unlike de Corte's implementation of Brogden classification, the empirical MNL optimal classification model is not dependent on approximately normally distributed criterion estimates. Classifying applicants directly using estimated attrition probabilities yielded more accurate and stronger evidence of incremental classification benefits of TAPAS over ASVAB for minimizing 6-month attrition.

Figure 1. **Comparison of 6-Month Attrition Probability Distributions Under Optimal Classifications Using Linear Predictor and Logistic Probability with 10% Rejection and Cut Score and Gender Restriction**



Hybrid Optimal Classification-Choice Model

We briefly describe a hybrid optimal classification model that includes applicant preferences. Using a utility interpretation, we can further modify the augmented criterion estimate to allow applicant preferences to affect the optimal classification solution. For example, letting Z_{ij} represent the incentive for the *j*th job that the *i*th applicant is eligible for, the augmented criterion estimate can be modified as follows so that classification optimization is conditional on applicant preferences for the incentives:

$$X^*_{(\alpha^r)ij} = \lambda\left(A_j + X_{ij}\right) + BZ_{ij} + E_{ij}$$

In the equation above,

B is a fixed parameter in the MNL model, with unknown constants A_js that will be computed to satisfy the job quota constraints for given scale parameter. Note that the additional term BZ_{ij} dilutes the effect of the original criterion estimate, X_{ij}, in identifying the job that will provide the greatest contribution to the average predicted criterion estimate of applicants. Instead of maximizing the criterion estimates X_{ij}s subject only to job quota constraints, the modified classification rule would classify an applicant to a job with lower predicted criterion estimate if the applicant's preference for the incentive offered is relatively higher.

We outline below a general approach for carrying out this hybrid model for solving an optimal classification problem conditional on applicant preferences for incentives. First, using data with actual applicant job choices and incentives Z_{ij}s, estimate a job choice model with utility equations

$$U_{ij} = A^*_j + \lambda^* X_{ij} + B^* Z_{ij} + E_{ij}$$

In the utility equation above, the term $\lambda^* X_{ij}$

represents the "observed classification effect" in the data. In previous studies, we found statistically significant classification effects for AA scores of applicants (Diaz, Ingerick, & Sticha, 2007a; Diaz, Ingerick, & Sticha, 2007b; Diaz, Sticha, Hogan, Mackin, & Greenston, 2012). Unlike in the classification problem, the constants A^*_js represent average preferences not explained by the incentive and classification composite. Next, expand the modified augmented criterion estimate as

$$X^*_{(\alpha^r)ij} = \lambda\left(A_j + X_{ij} + BZ_{ij}\right) + E_{ij}$$

with $B = B^*/\lambda$

, where B^* is the coefficient of the incentive estimated from the choice model. This expanded augmented criterion estimate is equivalent to the scaled utility equation for the MNL model, where B is known and fixed and A_js are job specific constants that satisfy the quota constraints. As before, these job specific constants can be solved using standard MNL estimation procedures.

Conceptually, the coefficient B in the expanded augmented criterion estimate represents the dilutive effect of incentives in the hybrid classification model. Setting the scale constant to λ^* will induce an optimal classification model in which the effect of applicant preferences for incentives relative to criterion estimates is equal to that observed in the data. This classification model will in fact be equal to the choice model except that the job specific constants are determined to satisfy target job quota constraints instead of the observed job choice proportions. We can use the scale constant to impose realistic limits to potential classification benefits or effectiveness of composites or predictors. Scale values larger than λ^* presume classification effects that are greater than observed. Setting the scale constant to extremely large enough values would effectively eliminate the effects of incentives in the classification. From a practical view, the important point is that competing policies, such as incentive programs, can potentially limit the classification benefits of alternative predictors for maximizing average performance. The scale constant provides a mechanism for imposing this limit if desired. Conversely, the hybrid model can also inform how alternative incentive policies affect average predicted performance of applicants.

Discussion and Recommendations

In this research we showed how the Brogden classification framework can be formulated using a DCM framework. We first specified a MMNL classification model that is analytically comparable to the MVN-based solution proposed by de Corte. To obtain the MMNL classification model, we rescaled the augmented criterion estimates proposed by Brogden using an arbitrarily large constant and then added independent standard Gumbel errors. This transformation produced modified augmented criterion estimates with the same structure as the utility equations in a MMNL model with error components corresponding to the original criterion estimates. Taking advantage of the similarity in interpreting job specific constants, we showed analytically and by numerical examples that the constants that satisfy the job quota constraints in Brogden's classification problem can be obtained using the standard MMNL model estimation method. We also showed how to specify a MMNL selection-classification model by interpreting the modified augmented criterion estimates as utility equations and applying random utility maximization.

We also proposed an empirical or sample based MNL classification model for evaluating potential classification efficiency that accommodates personnel classification policy constraints and is robust with regard to the form and distribution of criterion estimates. For classification problems where a large amount of applicant data is available, the empirical distribution of criterion estimates is a more accurate representation of the true distribution in the applicant population. The MNL classification model produced classification probabilities by fractionally weighting individual applicants to different jobs to obtain the highest overall predicted criterion.

We illustrated the sample data based MNL optimal classification model by evaluating the classification efficiency potential of ASVAB and TAPAS for minimizing 6-month attrition in four MOS under pure classification and selection-classification models and varying classification policy requirements using data from the TOPS study. While changes in estimated attrition due to inclusion of cut scores and gender restriction were negligible, their directions were consistent

with anticipated effects of the policy constraints. We expect these effects to be stronger with a more valid criterion model and greater differentiation in eligibility constraints among jobs. The 6-month attrition example showed that the choice of classification composites can substantially affect the results of optimal classification analysis. Using the linear predictor in the logistic model as a proxy for attrition probability estimates produced somewhat lower overall classification efficiency compared to direct classification using attrition probability estimates, and inaccurately indicated that TAPAS does not provide incremental classification benefit for MOS 11B beyond the ASVAB. The same inaccurate indication is expected from the MVN distribution based analytic approach of de Corte, using a linear predictor as the best MVN distributed proxy for the attrition probability estimate.

The constant λ used to scale the modified criterion estimates can have varying theoretical and practical implications in applications of the MNL classification model. On one extreme, as the scale approaches zero, the Gumbel error term will dominate the modified augmented criterion estimate, producing classification that would be close to random. On the other extreme, as the constant becomes large, the elements in the optimal classification probability matrix in the MNL classification model would generally approach ones or zeroes, assuming criterion estimates are continuous variables. This probability matrix of ones and zeroes is comparable to the decision matrix in a binary integer linear programming (BILP) model for matching applicants to jobs to maximize the average predicted criterion. In other words, the empirical MNL classification model converges to the BILP model as the scale constant increases to infinity. For values not large enough to produce classification probabilities of ones or zeroes, the scale constant has a practical interpretation as a parameter of "classification uncertainty." This uncertainty may be viewed as unreliability or measurement errors of criterion and predictor variables. It can also be related to shrinkage corrections to obtain cross-validated prediction results. These interpretations can be used to specify a scale constant that will produce the desired correction due to unreliability and/or shrinkage when evaluating the potential classification efficiency of predictors.

We also outlined a hybrid classification-choice model for evaluating the classification potential of predictors that accounts for applicant preferences. In real world classification, applicant preferences for incentives can potentially limit the classification benefits of alternative predictors. The hybrid approach provides a mechanism for imposing this limit if desired. Conversely, the hybrid model can also inform how alternative incentive policies affect average predicted performance of applicants.

In sum, the DCM provided an alternative framework for carrying out Brogden's optimal classification of applicants to jobs. The mathematical population based MMNL model specified in this research provides a computationally more convenient alternative to the MVN based method for evaluating potential classification efficiency. The sample data based MNL model is more robust with respect to distributional assumptions about the criterion estimates and can easily accommodate policy constraints, thereby producing more accurate results with respect to potential classification efficiency of alternative predictors. Lastly, concepts underlying the DCM based framework for classification can be applied to develop a hybrid classification-choice model that can be used to add realism to optimal classification analysis (e.g., by considering the effects of enlistment incentives and applicant job preferences).

References

Alley, W. E., & Darby, M. M. (1994). Estimating the benefits of personnel selection and classification: an extension of the Brogden table. *Educational and Psychological Measurement, 55,* 938-958.

Ben-Akiva, M. & Lerman (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: MIT Press.

Bierle, M. (2003). An introduction to BIOGEME (Version 1.3) http://roso.epfl.ch/biogeme.

Brogden, H. E. (1954). A simple proof of a personnel classification theorem. *Psychometrika, 19,* 205-208.

Brogden, H. E. (1955). Least square estimates and optimal classification. *Psychometrika, 20,* 249-252.

Brogden, H. E. (1959). Efficiency of classification as a function of number of jobs, percent rejected, and the validity and intercorrelation of the job performance estimates. *Educational and Psychological Measurement, 19*, 181-190.

Cheng, C., & Darby, M. M. (1997). Efficiency of classification: a revision of the Brogden table. (AL/HR-TR-1997-0013). Brooks AFB, TX: Manpower and Personnel Research Division, Armstrong Laboratory, Human Resources Directorate.

DeCorte, W. (2000). Estimating the classification efficiency of a test battery. *Educational and Psychological Measurement, 60,* 73-85.

Diaz, T., Ingerick, M., & Sticha, P. (2007a). *Modeling Army applicant's job choices: The EPAS Simulation Job Choice Model (JCM)* (Study Note 2007-01). Arlington, VA: U.S. Army Research Institute for the Behavior and Social Sciences.

Diaz, T., Ingerick, M., & Sticha, P. (2007b). *Raising the enlistment bonus cap: forecasted impact on Army accessions* (Study Report 2007-02). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Diaz, T. E., Sticha, P. J., Hogan, P, Mackin, P, & Greenston, P (2012). *Determinants of the Army Applicant Job Choice Decision and the Development of a Decision Support Tool for the Enlistment Incentive Review Board* (Technical Report 1301). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Ingerick, M., Diaz, T., & Putka, D. (2009). *Investigations into Army enlisted classification systems: Concurrent validation report* (Technical Report 1244). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Knapp, D. J. & Heffner, T. S. (Eds.) (2012). *Tier one performance screen initial operational test and evaluation: 2011 interim report* (Technical Report 1306). Arlington, VA: U.S. Army Research Institute for the Behavior and Social Sciences.

McKelvey, R., and Zavoina, W. (1975). A Statistical Model for the Analysis of Ordinal Level Dependent Variables. *Journal of Mathematical Sociology 4*, 103-120.

Sticha, P.J., Diaz, T.E., Greenston, P.M., & McWhite, P.B. (2007). *Field evaluation of enlisted personnel allocation system (EPAS) enhancements to the recruit quota system (REQUEST)* (Technical Report 1212). Arlington, VA: U.S. Army Research Institute for the Behavior and Social Sciences.

Train, K. (1986). *Qualitative choice analysis*. Cambridge, MA: MIT Press.

Train, K. (2003). *Discrete choice methods with simulation*. New York, NY: Cambridge University Press.

Trippe, D. M., Diaz, T. E., & Ingerick, M. (2012). Evaluation of TAPAS potential for classification purpose. In D. J. Knapp & T. S. Heffner (Eds.), *Tier one performance screen initial operational test and evaluation: 2011 interim report* (Technical Report 1306). Arlington, VA: U.S. Army Research Institute for the Behavior and Social Sciences.

Appendix A: Derivations Under Modified Augmented Criterion

Job Quota Constraints Equation

We begin by specifying $f^*_{K_i^h}(k)$, the probability that the *k*th job of a randomly selected applicant has the highest modified augmented criterion under the modified optimal classification algorithm. Using $X_{(a^r)ij}$s for classifying applicants, the optimal job is determined jointly by $\boldsymbol{X}_i = \langle X_{i1}, \ldots, X_{im} \rangle$ and $\boldsymbol{E}_i = \langle E_{i1}, \ldots, E_{im} \rangle$. We use conditional probability in our derivation below to break the joint probability of $\langle \boldsymbol{X}_i, \boldsymbol{E}_i \rangle$ into manageable parts.

Conditioning on $\boldsymbol{X}_i = \boldsymbol{x}_i$ and integrating over all possible values $\boldsymbol{x}_i$, the probability that the *k*th job has the highest modified augmented criterion estimate is obtained as follows:

$$\begin{aligned} f^*_{K_i^h}(k) &= P\big(\lambda(A_j + X_{ij}) + E_{ij} < \lambda(A_j + X_{ik}) + E_{ik}\ ;\ j \neq k\big) \\ &= \int_{x_i \in R^m} P\big(\lambda(A_j + x_{ij}) + E_{ij} < \lambda(A_k + x_{ik}) + E_{ik} \big| \boldsymbol{X}_i = \boldsymbol{x}_i\ ;\ j \neq k\big) f_{X_i}(\boldsymbol{x}_i) d\boldsymbol{x}_i \\ &= \int_{x_i \in R^m} P\big(K_i^h = k \big| \boldsymbol{X}_i = \boldsymbol{x}_i\big) f_{X_i}(\boldsymbol{x}_i) d\boldsymbol{x}_i \end{aligned}$$

Substituting the MNL conditional probability expression for $P\big(K_i^h = k \big| \boldsymbol{X}_i = \boldsymbol{x}_i\big)$, we obtain the system of nonlinear equations below, with variables A_js, representing job quota constraints under the modified classification algorithm:

$$\int_{x_i \in R^m} \frac{\exp\big(\lambda(A_k + x_{ik})\big)}{\sum_{j=1}^m \exp\Big(\lambda\big(A_j + x_{ij}\big)\Big)} f_{X_i}(\boldsymbol{x}_i) d\boldsymbol{x}_i = q_k\ , k = 1, \ldots, m$$

MPP Equation

We derive $f^*_{X_i^h}(x)$, the PDF of the estimated criterion score corresponding to the optimal job of a randomly chosen applicant under the modified classification algorithm. To begin, we note that conditional on observing the *m* criterion estimates,

$$\begin{aligned} f^*_{X_i^h}(x | \boldsymbol{X}_i = \boldsymbol{x}_i) &= \textstyle\sum_{k=1}^m P\big(K_i^h = x_k \big| \boldsymbol{X}_i = \boldsymbol{x}_i\big) 1_{(X_k = x)}\big) \\ &= \textstyle\sum_{k=1}^m P\big(K_i^h = k \big| \boldsymbol{X}_i = \boldsymbol{x}_i\big) 1_{(X_k = x)}\big) \end{aligned}$$

Note that $P\big(K_i^h = k \big| \boldsymbol{X}_i = \boldsymbol{x}_i\big)$ is the MNL probability defined earlier and $1_{(X_k = x)}$ is an indicator function that evaluates to one if $X_k = x$ and zero otherwise. The first line above simply means that the PDF of X_i^h given $\boldsymbol{X}_i = \boldsymbol{x}_i$ is the probability that the optimal predicted criterion score can be anyone of the *m* observed predicted criterion scores**.** The indicator function $1_{(X_k = x)}$ ensures that $f^*_{X_i^h}(x | \boldsymbol{X}_i = \boldsymbol{x}_i)$

evaluates to non-zero probabilities only at the observed criterion estimates (i.e., $x_{i1}, \ldots, x_{im}$). The second line is obtained by noting that $X_i^h = x_k$ implies $K_i^h = k$.

Using the above conditional probability representation, we derived below the unconditional PDF of X_i^h. To facilitate the derivation below, we partition the vector of criterion estimates into the scalar X_k, representing the criterion estimate for the *k*th job, and the (*m*-1) vector $\boldsymbol{X}_{(k)}$, representing all other criterion estimates. Again, by conditioning on $\boldsymbol{X}_i = \boldsymbol{x}_i$ and integrating over all possible values $\boldsymbol{x}_i$, we can derive $f^*_{X_i^h}(x)$ as follows:

$$
\begin{aligned}
f^*_{X_i^h}(x) &= \int_{x_i \in R^m} f^*_{X_i^h}(x|\boldsymbol{X}_i = \boldsymbol{x}_i)\, f_{\boldsymbol{X}_i}(\boldsymbol{x}_i) d\boldsymbol{x}_i \\
&= \sum_{k=1}^{m} \int_{x_i \in R^m} P\left(K_i^h = k \middle| \boldsymbol{X}_i = \boldsymbol{x}_i\right) 1_{(X_k = x)}) f_{X_k}(x_k) f_{\boldsymbol{X}_{(k)}|X_k}\left(\boldsymbol{x}_{(k)}|x_k\right) d\boldsymbol{x}_i \\
&= \sum_{k=1}^{m} \int_{x_i \in R^{m-1}} P\left(K_i^h = k \middle| \boldsymbol{X}_{(k)} = \boldsymbol{x}_{(k)}, X_k = x\right) f_{X_k}(x) f_{\boldsymbol{X}_{(k)}|X_k}\left(\boldsymbol{x}_{(k)}|x\right) d\boldsymbol{x}_{(k)} \\
&= \sum_{k=1}^{m} f_{X_k}(x) \int_{x_i \in R^{m-1}} \frac{\exp\left(\lambda(A_k + x)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_{\boldsymbol{X}_{(k)}|X_k}\left(\boldsymbol{x}_{(k)}|x\right) d\boldsymbol{x}_{(k)}
\end{aligned}
$$

Again,this PDF involves the unknown job specific constants as parameters. To completely specify $f^*_{X_i^h}(x)$ we need to solve the job quota system of nonlinear equations for A_js.

After completely specifying $f^*_{X_i^h}(x)$ we can evaluate the MPP under the modified optimal classification algorithm using

$$
\begin{aligned}
MPP &= \sum \int_{R^1} x f_{X_k}(x) \int_{x_i \in R^{m-1}} \frac{\exp\left(\lambda(A_k + x)\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_{\boldsymbol{X}_{(k)}|X_k}\left(\boldsymbol{x}_{(k)}|x\right) d\boldsymbol{x}_{(k)}\, dx \\
&= \sum_{k=1}^{m} \int_{x_i \in R^m} x \frac{\exp\left(\lambda(A_k + x_{ik})\right)}{\sum_{j=1}^{m} \exp\left(\lambda\left(A_j + x_{ij}\right)\right)} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}
\end{aligned}
$$

Appendix B: Biogeme MMNL Model Approximation Examples

The following examples describe the utility equations and input data requirements of an MMNL model for approximating Brogden optimal classification using the Biogeme model file syntax and data format. In these examples we will evaluate the classification efficiency of a predictor battery for four jobs. We will assume that the predicted criterion estimates have been standardized and that we have available the covariance (correlation) across the four jobs. The three examples below demonstrate the MMNL approximation for Brogden optimal classification depending on the structure of the predictor covariance/correlation and type of problem, whether classification-only or simultaneous selection-classification.

Example 1: Classification with Uncorrelated Predictors

This example shows how to specify an MMNL model and input data to approximate Brogden's optimal classification when predictors are uncorrelated and the accession rate is 100 percent (i.e., pure classification problem). For this problem, assume that the predicted criteria are uncorrelated across jobs with squared-validities equal to 0.50, 0.60, 0.70, and 0.50 and the target job allocations are 40%, 30%, 20%, and 10%.

The top part in Figure 1 shows sections of the Biogeme model file specifying an MMNL model for approximating an optimal classification problem with four jobs and uncorrelated predictors. The section labeled [Utilities] specifies the utility equations for the four jobs (alternatives), with each row representing one job. The first three columns provide, respectively, unique ID numbers for identifying the "chosen alternative" in the input data file, unique alphanumeric labels for identifying alternatives in Biogeme's output, and the name of the variable in the input data file that indicates "availability" of each alternative. The last column (shown in bold text) describes the utility equations using a parameter×variable syntax. Using this syntax, the term " " represents the job-specific constants for the first job, where is the (fixed) job specific constant parameter and " " is a variable whose constant value (equals 1) is specified under the [Expressions] section of the model file. The criterion estimates are represented by the error component terms using the syntax " " where the string " " represents a random parameter with mean and standard deviation .

Figure B-1. **Model File and Input Data for Approximating a Classification Problem with Uncorrelated Predictors**

Selected Sections of the Model File:

Input Data:

The section labeled [Beta] describes how parameters are handled during estimation. The first column identifies the parameters in the utility equations. The next three columns specify starting values and bounds for the parameters, while the last column indicate whether the value of a parameter will be fixed at the specified starting value (status=1) or will be estimated in the problem (status=0). For approximating Brogden's optimal classification, the only parameters in the MMNL model to be estimated are the job specific constants (i.e., ASC).[6] The means and variances of the criterion estimates (error components in the utility equation) are usually computed separately from analysis data. For this example, the predictors have means all equal to zero and standard deviation 0.707107, 0.774597, 0.836660, and 0.707107 (i.e., validities corresponding to R^2 values of 0.50, 0.60, 0.70, and 0.50).

The last two sections in the model file, [Group] and [Scale], are used to specify the arbitrary value of the scaling parameter λ. In normal applications of Biogeme, these two sections are used in combination to specify a scale heterogeneous model (i.e., with different scales across groups of observations or applicants). For our particular problem, we use these sections to specify an arbitrary scale for the entire data, as follows. The expression (
) under the [Group] section evaluates to 1 for all four observations/jobs (see following discussion), effectively specifying a single scale for all four jobs. This scale is specified as a fixed parameter under the [Scale] section. For this example we are using $\lambda = 100$.

The bottom part in Figure 1 shows the desired format of the input data for the MMNL model approximation in this example. The first row identifies the variables corresponding to the columns in the data. The data shown in Figure 1 is in weighted form, with each row observation representing a portion of the estimation data. Instead of showing individual observations (as in a JCM estimation problem), this data format describes the distribution of applicants who "chose" (or will be assigned to) each job. The "chosen" job represented in each row is identified under the column, while the percentage of applicants who "chose" each job is given under the column. To obtain estimated ASCs that satisfy the job quota constraints, we specify weights that are equal or proportional to the quota percentages of the job corresponding to each row. In this example, the job quota constraints are 40%, 30%, 20%, and 10%, respectively, for jobs 1, 2, 3, and 4. Therefore, we created a choice data with applicant job choice percentages equal to these values.

Example 2: Classification with Correlated Predictors

This example shows how to specify an MMNL model and input data to approximate Brogden's optimal classification when predictors are correlated and the accession rate is 100 percent (i.e., pure classification problem). We will use the same set of job quota constraints, 40%, 30%, 20%, and 10%, as in the first example, but specify the job predictor covariance

[6] Note that normalization involves fixing the value of one of the ASCs. In this example we fixed the ASC for the first job at the final estimated value obtained using our implementation of De Corte's method. Our implementation of De Corte's method did not set the constant for the first job to zero, as he recommended. In order to make direct comparison between the constants obtained under the two methods, we fixed the ASC in the approximating MMNL model to that obtained using the multivariate normal approach.

$$\mathbf{V} = \begin{bmatrix} 0.5 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.6 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.7 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.5 \end{bmatrix}$$

Figure 2 shows the required Biogeme MMNL model file when criterion estimates are correlated. Because we are using the same job quota constraints as in the first example, the required input data will be the same as that shown in Figure 1. The utility equations specified in Figure 2 involve job-specific constants and random error components with longer expressions, which are based on the Cholesky decomposition of the covariance matrix $\mathbf{V} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$ where $\mathbf{L}$ is a lower triangular matrix with positive diagonal entries. This factorization is often used to construct or generate a random vector with specified covariance from independently distributed random variables. For example, if $\boldsymbol{Z} = (Z_1, \ldots, Z_m)^T$ is a random vector with elements that are independently distributed as standard normal, then $\boldsymbol{F} = \mathbf{L}\boldsymbol{Z}$ is a random vector that is distributed as multivariate normal with zero mean and variance $\mathbf{V}$.

In Figure 2 the error components in the utility equations are just the rows in the 4-dimensional vector $\boldsymbol{F}$ obtained below:

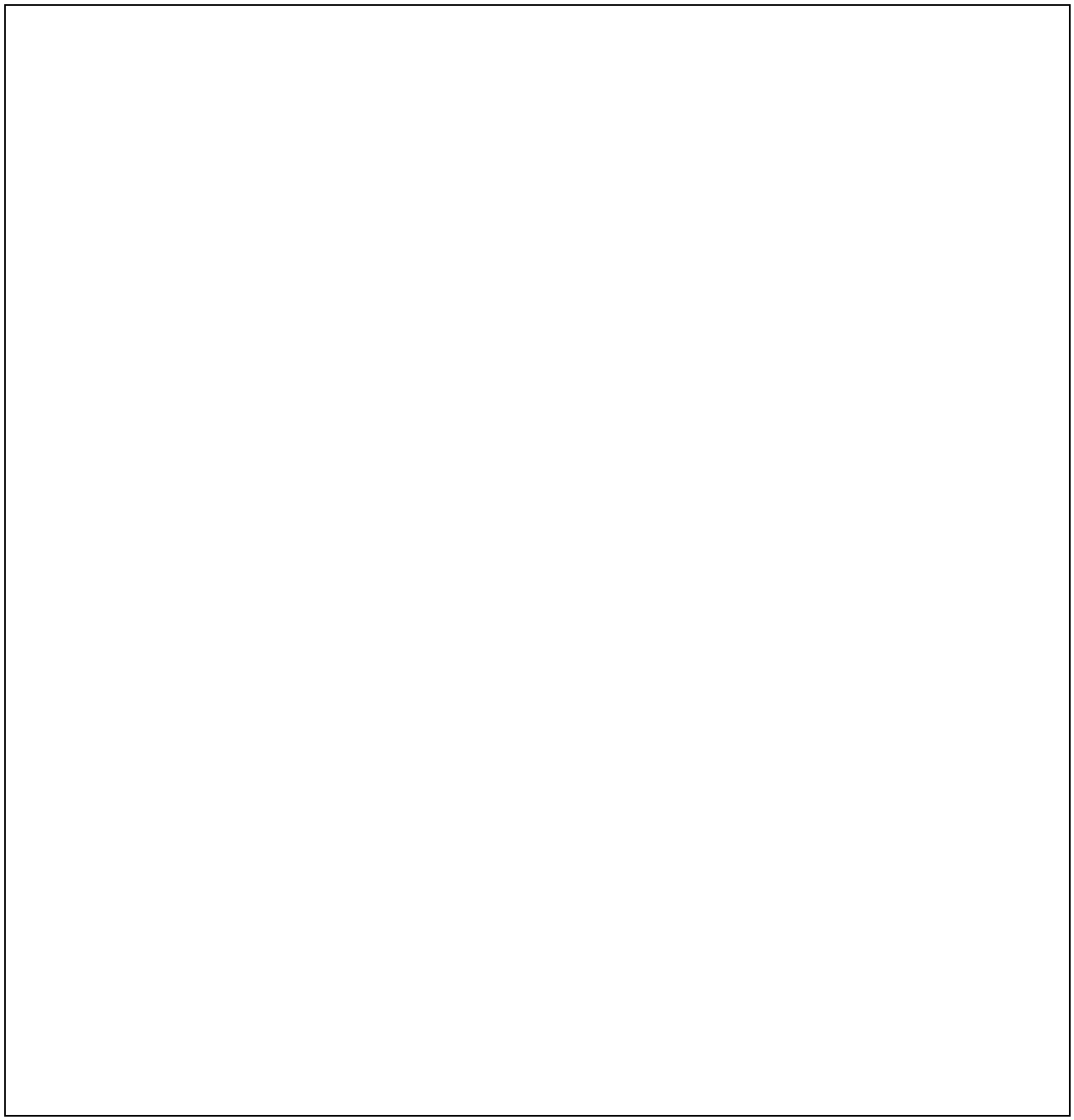
$$\boldsymbol{F} = \mathbf{L}\boldsymbol{Z}$$

$$= \begin{bmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ L_{31} & L_{32} & L_{33} & 0 \\ L_{41} & L_{42} & L_{43} & L_{44} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{bmatrix}$$

$$= \begin{bmatrix} Z_1 L_{11} \\ Z_1 L_{21} + Z_2 L_{22} \\ Z_1 L_{31} + Z_2 L_{32} + Z_3 L_{33} \\ Z_1 L_{41} + Z_2 L_{42} + Z_3 L_{43} + Z_4 L_{44} \end{bmatrix}$$

The values of the elements of the lower triangular matrix L_{ij}s) are specified under the section [Expressions] as variables with fixed values across observations. The standard normal random variables Z_is are specified using Biogeme's random parameter syntax (e.g., $Z_1 =$ Fmu0 [Fsd1], $Z_2 =$ Fmu0 [Fsd2], etc).

Figure B-2. **Selected Sections of the Model File for Approximating a Classification Problem with Correlated Predictors**



Example 3: Selection-Classification with Correlated Predictors

This example shows how to specify an MMNL model and input data to approximate Brogden's optimal classification when predictors are correlated and the accession rate is less than 100 percent (i.e., selection-classification problem). For this example we will use a rejection rate of 30% and distribute the remaining 70% using the percentages in the first two examples. This allocation produces job quota constraints of 28%, 21%, 14%, and 7% for the four jobs and 30% for the dummy job for non-accession. We also specify criterion predictor correlations using the covariance matrix given in the second example.

Figure 3 shows the [Beta] and [Utilities] sections in the Biogeme model file and the input data needed for solving the selection-classification problem using an MMNL model. In specifying the Biogeme model file, we simply modified the model file used in the second example by adding the utility equation for the non-accession job. The input data is also modified by adding a fifth observation with weight equal to the rejection rate and setting the weights for the other four observations to 28%, 21%, 14%, and 7%. All other components of the model file are the same as in Figure 2.

Figure B-3. **Selected Sections of the Model File and Input Data for Approximating a Selection-Classification Problem with Correlated Predictors**

Selected Sections of the Model File:

Input Data:

Example 4: Empirical (Sample Based) MNL Classification Model

Figure 4 shows relevant sections of the Biogeme model file and the structure of the input data for carrying out the MNL classification model analysis using criterion estimates that were simulated from a multivariate normal distribution with zero mean and the same covariance matrix as in the Example 2. There are two main changes in Figure 4 compared to the MMNL model files in Example 2. First, individual applicant criterion estimates observations (X1, X2, X3, X4) are directly included in the model, instead of being represented as error components. Second, actual sample applicant observations comprised the input data, with each observation replicated as many times as the number of jobs using job quota percentages as replicate weights. Technically, the sample observations are used as "grid points" for the empirical distribution

represented by the sample data.[7] Note that correlation of criterion estimates are implicitly taken care of by the sample data. As before, we use a large value for the scale parameter ($\lambda = 100$) to ensure that differences between criterion estimates drive the assignment algorithm.

[7] Under the simulated MLE algorithm for the MMNL approximation method, the grid points correspond to the random values generated to approximate the multivariate normal distribution of the error components/criterion estimates.

Figure B-4. **Selected Sections of the Biogeme Model File and Input Data for the MNL Classification Efficiency Analysis Model**

Selected Sections of the Model File:

Input Data: